

PERANCANGAN DAN IMPLEMENTASI SISTEM ROBOT KURSI RODA MENGUNAKAN SPEECH RECOGNITION

DESIGN AND IMPLEMENTATION OF SPEECH RECOGNITION BASED WHEEL CHAIR ROBOT SYSTEM

Hendriadi Mukri¹ , Inung Wijayanto,S.T.,M.T.² , Unang Sunarya,S.T.,M.T.³

Prodi S1 Teknik Telekomunikasi, Fakultas Teknik Elektro, Universitas Telkom

¹mukri.hendriadi@gmail.com, ²twijayanto@telkomuniversity.ac.id,

³unangsunarya@telkomuniversity.ac.id

Abstrak

Di Indonesia, masih banyak penduduk yang mengalami kecacatan fisik seperti kelumpuhan yang menyebabkan mereka tidak dapat bergerak dengan sendirinya. Contoh pada tahun 2009 menurut Badan Kemntrian Sosial RI jumlah penyandang cacat pada sembilan provinsi sebanyak 299.203 jiwa dan 10,5% (31.327 jiwa) merupakan penyandang cacat berat yang mengalami hambatan dalam kegiatan sehari-hari.

Kursi roda otomatis merupakan alat yang difungsikan untuk meringankan beban penderita kelumpuhan tangan dan kaki yang tidak dapat menggerakkan kursi roda manual pada umumnya. Prinsip kerja dari alat ini adalah dengan memanfaatkan gelombang suara pada manusia sebagai masukan informasi pada alat untuk dapat menentukan arah gerak dari kursi roda otomatis tersebut. Gelombang suara manusia yang masuk ke Raspberry Pi di ekstraksi menggunakan metode MFCC. Hasil ekstraksi tersebut dibandingkan dengan data ekstraksi yang telah ada sebelumnya pada *database* dengan menggunakan metode KNN. Metode KNN sendiri berkerja dengan cara menghitung jarak terdekat suatu sinyal baru dengan sinyal yang sudah teridentifikasi sebelumnya, jumlah sinyal yang dibandingkan tergantung dengan nilai K yang dimasukkan, adapun untuk menghitung jarak antar sinyal, KNN menggunakan metode *Euclidean distance*. Hasil dari perbandingan ekstraksi dengan menggunakan metode KNN, didapat beberapa sinyal yang ada pada *database* mempunyai kemiripan dengan sinyal masukkan yang baru. Sehingga setelah dibandingkan didapat hasil berupa keputusan perintah yang sesuai dengan *database* untuk menggerakkan motor dari kursi roda.

Penelitian pada tugas akhir ini telah berhasil membuat sistem yang dapat mengenali ucapan dengan waktu pengambilan keputusan adalah 5.59 detik pada saat kondisi sepi dan jarak mikrofon terhadap pusat suara adalah 10 cm, namun nilai yang didapatkan tersebut hanya pada saat pengujian tanpa kursi roda, sedangkan untuk pengujian dengan kursi roda waktu pengenalan ucapannya menjadi 14.9 detik.

Kata Kunci : Robot kursi roda, MFCC, KNN, *Euclidean Distance*

Abstract

In Indonesia, there's still many its people has physical disability such as paralysis that make them cant move by them self. For example on 2009 amount of physical disability in 9 Provinces is 299.203 soul and 10.5% (31.327 soul) disability severe who has limitation of Undergo daily activity (Ministry of Social Affairs RI).

Automatic Wheelchair is technology machine for lighten up burden of hand or feet paralysis who cant drive the wheelchair manually. Principal work of this machine is using sound wave of human/user as information input for determine direction drive of wheelchair automatically Sound wave of human will be process in raspberry pi and will extraction using MFCC method, the result of the extraction will be compared with extraction data which has been exist before by using KNN method. This KNN method work by counting nearest distance the new signal with the signal which has been identification before. Amount of signal will be compared, depends on value of K as input. And for calculate distance between signal, KNN using *Euclidean distance* methode. The result of comparing extration by using KNN methode will be obtained several signal which exist on *database* has similar new signal input. So after compared, will got result decision or command according to the *database* that will drive the motor from the wheelchair.

Research in this final project has succeed to make speech recognition system to drive the wheelchair with accuracy 100 % in quiet condition and the distance of microphone from source is 10 cm with the decision time is 4.64 second.

Keyword : wheelchair robot, MFCC, KNN, *Euclidean Distance*

1. Pendahuluan

Dewasa ini teknologi sudah berkembang dengan pesat, khususnya bidang robotika. Bidang robotika sudah banyak membawa dampak positif bagi kehidupan manusia. Dampak positif dari berkembangnya bidang robotika dapat dilihat dalam berbagai bidang kehidupan diantaranya bidang kesehatan, bidang industri dan banyak bidang lainnya. Contoh pada bidang industri yakni *arm* robot yang biasa difungsikan untuk menggantikan tugas manusia dalam hal produksi suatu produk perusahaan. Adanya *arm* robot pada bidang industri, mengakibatkan meningkatnya kualitas produksi karena dapat mengurangi *human error* yang biasa terjadi pada proses produksi.

Dalam bidang kesehatan, robot banyak digunakan untuk membantu orang yang sedang sakit seperti robot kursi roda yang dapat membantu orang yang lumpuh bergerak tanpa harus ada dorongan dari orang itu sendiri. Robot kursi roda sendiri ada beberapa jenis berdasarkan jenis masukannya, diantaranya robot kursi roda berdasarkan masukan dari tombol, suara dan dari otak. Pada robot kursi roda dengan masukan suara, robot kursi roda memanfaatkan masukan suara dari pengguna kursi roda untuk diolah dalam suatu sistem sehingga menghasilkan keluaran berupa perintah terhadap kursi roda untuk bergerak sesuai dengan perintah dari penggunaanya.

Pada perkembangannya, robot kursi roda dengan masukan suara yang telah ada saat ini adalah robot kursi roda yang dibuat dengan memanfaatkan *Google Voice* sebagai pengolah suara guna dimanfaatkan untuk mengatur arah gerak robot. Untuk mengatasi permasalahan ketersediaan jaringan internet yang belum menjangkau semua wilayah di Indonesia, diperlukan robot kursi roda yang dapat digunakan tanpa adanya internet (*offline*).

Pada tugas akhir ini penulis membuat sebuah sistem pengenalan suara manusia untuk mengendalikan sebuah robot kursi roda. Sistem pengenalan suara dirancang menggunakan *library CMU Sphinx*, dimana *CMU Sphinx* menggunakan metode ekstraksi ciri MFCC dengan jumlah filter bank 40, jumlah koefisien MFCC 13 dan FFT sebanyak 512.

2. Dasar Teori

2.1 Robot

Robot merupakan suatu alat yang dirancang secara khusus oleh manusia yang didalamnya terdapat *processor* dan *controller* yang difungsikan untuk tugas tertentu baik dikontrol langsung oleh manusia maupun ditanamkan kecerdasan buatan untuk melakukan fungsinya secara otomatis^[1].

2.2 Robot Kursi Roda

Robot kursi roda merupakan robot yang berbentuk kursi roda yang dapat bergerak berdasarkan masukan dari penggunaanya. Robot ini dapat menjadi alat bantu oleh penderita kelumpuhan tangan dan kaki yang mengalami kesulitan berjalan baik dikarenakan oleh penyakit, cedera, atau cacat. Masukan yang digunakan pada robot kursi roda dapat berupa *remote control*, suara, maupun sinyal otak. Adapun fokus pembahasan dari tugas akhir ini yakni robot kursi roda berdasarkan masukan suara sebagai *controller* nya.

2.3 Single Board Computer

Single Board Computer (SBC) merupakan perangkat komputer yang dikemas dalam sebuah PCB yang didalamnya terdapat *microprocessor*, memori serta antarmuka I/O. Idealnya sebuah SBC mempunyai kemampuan sama dengan sebuah komputer pada umumnya hanya saja ukuran dari SBC jauh lebih kecil dari komputer pada umumnya. Awalnya SBC hanya difungsikan untuk pembelajaran, pemodelan sistem, serta kontrol khusus. Namun belakangan ini SBC difungsikan untuk aplikasi yang lebih rumit. *Raspberry pi* merupakan salah satu contoh SBC yang cukup terkenal dikalangan mahasiswa/pelajar dikarenakan *Raspberry pi* yang mempunyai harga yang relatif murah serta spesifikasi yang cukup memadai^[1]. Adapun pada Tugas akhir ini penulis menggunakan *Raspberry Pi 3* sebagai SBC nya dikarenakan sesuai dengan kebutuhan *speech recognition*.

2.4 Mikrokontroler

Mikrokontroler merupakan salah satu peranti kontrol yang digunakan sebagai kendali dari sistem tertanam. Sebagai peranti kontrol, ia memiliki sifat yang dapat diprogram oleh pemakai.

Adapun pada Tugas Akhir ini penulis menggunakan *Arduino* sebagai mikrokontrolernya.

2.5 Automatic Speech Recognition

Automatic speech recognition (ASR) merupakan adalah suatu sistem yang berfungsi menerjemahkan suatu bahasa lisan menjadi suatu bentuk data dalam komputer. Prinsip kerja ASR adalah dengan menyaring informasi yang didapat dari perangkat audio seperti mikrofon kemudian membandingkannya dengan *database* yang tersedia.

Secara konseptual bekerja dengan mengubah sinyal suara menjadi angka-angka dan membandingkan dengan kode yang terdapat dalam database untuk menerjemahkan masukan suara tersebut^[5].

Ada empat proses dalam sistem ASR, yaitu:

1. Tahap Menerima Masukan : pada tahap ini sistem mendapat masukan berupa gelombang suara. Suara ini dapat berasal dari tangkapan mikrofon ataupun hasil rekaman suara.
2. Tahap Ekstraksi : Pada tahapan ini sistem akan melakukan proses ekstraksi ciri dari sinyal yang didapat. Ada beberapa metode yang dapat dilakukan untuk ekstraksi ciri, misalnya Mel-Frequency Cepstral Coeficient (MFCC), Linear Predictive Coding (LPC) dan sebagainya.
3. Tahap Perbandingan : Pada tahap ini sistem akan membandingkan hasil ekstraksi yang sudah dilakukan dengan database yang tersedia.
4. Tahap Validasi : Pada tahap inilah sistem akan mengambil keputusan apakah suara masukan dapat dikenali atau tidak.

2.6 MFCC

MFCC merupakan salah satu jenis metode ekstraksi ciri yang digunakan sebagai vektor ciri yang baik untuk merepresentasikan suara manusia dan sinyal musik. Analisis suara pada *Mel-Frequency* didasarkan pada persepsi pendengaran manusia, karena telinga manusia telah diamati dapat berfungsi sebagai filter pada frekuensi tertentu. MFCC digambarkan dalam skala mel-frequency yang merupakan frekuensi linier dibawah 1000Hz dan logaritmik diatas 1000Hz.

Adapun beberapa tahap MFCC yakni :

2.6.1 Frame Blocking

Sinyal suara manusia merupakan sinyal yang tidak stabil. Namun dapat diasumsikan sebagai sinyal yang stabil pada skala waktu 10-30ms. Oleh karena itu dibutuhkan *frame blocking* untuk memotong sinyal menjadi bagian yang lebih kecil sehingga didapat karakteristik suara yang stabil. Pada *CMU Sphinx* sinyal suara dipotong setiap 25ms.

2.6.2 Windowing

Proses windowing dilakukan untuk mengurangi kebocoran spektral atau aliasing yang merupakan efek dari timbulnya sinyal baru yang memiliki frekuensi yang berbeda dari frekuensi aslinya. Hal tersebut dapat terjadi karena rendahnya sampling rate atau karena proses framing yang menyebabkan sinyal menjadi diskontinu. Pada *CMU Sphinx* besar hamming window yang digunakan adalah 0.025625.

2.6.3 FFT

Untuk mendapatkan sinyal dalam domain frekuensi salah satu metode yang dapat digunakan adalah DFT (Discrete Fourier Transform). DFT dilakukan terhadap semua frame yang telah di-windowing sebelumnya. Untuk mempercepat proses DFT dan mengurangi beban komputasi maka dilakukanlah FFT (Fast Fourier Transform) yaitu metode transformasi Fourier dengan lebih cepat. Pada *CMU Sphinx* ukuran FFT yang digunakan adalah 512. Berikut persamaan dari *Fast Fourier Transform (FFT)* :

$$F(w) = \sum_{n=0}^{N-1} Y(n) e^{-j2\frac{\pi}{N}wn} \quad (2.1)$$

Keterangan :

N = jumlah *frame* pada suatu data sinyal suara

Y(n) = sinyal input *frame* ke-n pada tahap ini

W = panjang DFT, dimana $0 < w < N-1$

F(w) = Hasil FFT pada DFT ke-w

2.6.4 Mel-Frequency Wrapping

Pada proses ini dilakukan perubahan skala frekuensi menjadi skala Mel. Skala Mel-Frequency adalah frekuensi linier dibawah 1KHz dan Logaritmik diatas 1KHz. Pada *CMU Sphinx* nilai filter yang digunakan adalah 40. Skala Mel dapat diperoleh dengan persamaan :

$$B(f) = 2595 \times \log\left(1 + \frac{f}{700}\right) \quad (2.2)$$

Keterangan :

B = skala Mel-Frequency

f = frekuensi linier

2.6.5 Cepstrum

Selanjutnya dilakukan DCT (Discrete Cosine Transform) untuk memperoleh sinyal dalam domain frekuensi. Hasil keluaran dari proses DCT ini disebut Mel-Frequency Cepstral Coefficient (MFCC). Pada sistem pengenalan suara biasanya hanya diambil 12 atau 13 koefisien pertama dari MFCC untuk mendapatkan hasil ekstraksi yang baik. Pada *CMU Sphinx* nilai koefisien MFCC yang digunakan adalah 13.

Koefisien MFCC diperoleh dengan persamaan :

$$y_t(k) = \sum_{m=1}^M \log(|y_t(m)|) \cos\left(k(m - 0,05)\frac{\pi}{M}\right), \quad k = 0, \dots, j \quad (2.3)$$

Keterangan :

$y_t(m)$ = frekuensi mel ke-m

M = jumlah frame

M = indeks frame

2.7 KNN

KNN adalah salah satu metode klasifikasi data. Prinsip dasar KNN adalah dengan menghitung jarak terdekat suatu sinyal baru dengan sinyal yang sudah teridentifikasi sebelumnya. Jarak terdekat sebuah sinyal baru akan digunakan untuk menentukan sinyal tersebut masuk kedalam kelas tertentu. Nilai k pada KNN menentukan seberapa banyak jarak terdekat yang dibandingkan. Misalkan nilai k=3, maka 3 tetangga terdekat akan dijadikan pembanding untuk menentukan kelas data baru tersebut. Dari ketiga data tersebut kemudian divoting kelas terbanyak dalam kumpulan data tersebut untuk selanjutnya dijadikan kelas bagi data yang baru di uji. Misal jika ada dua dari 3 data tersebut memiliki kelas yang sama maka data baru tersebut akan dianggap satu kelas dengan kedua data yang sama tersebut. Demikian seterusnya sesuai nilai k yang ditentukan. Penentuan jarak data pada KNN menggunakan metode euclidean distance. Untuk selengkapnya akan dijelaskan berikut ini^[3].

2.8 Euclidience Distance

Euclidean distance adalah perhitungan jarak dua buah titik yang berada dalam ruang Euclidean. Metode Euclidean adalah metode yang digunakan untuk menghitung jarak kedua buah titik tersebut. Pada dasarnya metode ini hanya digunakan untuk mengukur jarak dua buah titik yang berada pada bidang yang sama. Namun seiring perkembangan ilmu pengetahuan, metode ini telah dikembangkan dan dapat mengukur perbedaan pola dari suatu data. Atau dengan kata lain dapat dikatakan metode ini dapat mengukur kemiripan suatu data dengan data yang lain^[5].

3. Perancangan Sistem

Pada tugas akhir ini dirancang sebuah sistem pengenalan suara manusia yang difungsikan untuk menggerakkan robot kursi roda. Perintah-perintah untuk menggerakkan robot kursi roda sendiri sudah dimasukkan dan disimpan sebelumnya dalam sebuah *data base*. Perintah yang dimasukkan dibatasi hanya beberapa perintah yaitu *forward*, *turn left*, *turn right*, *stop*, dan *shutdown now*.

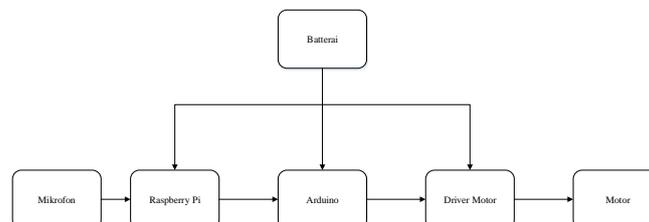
Masukan suara manusia ditangkap menggunakan mikrofon yang terhubung dengan *Raspberry Pi*. Pengguna robot kursi roda hanya memasukan kata yang telah disimpan pada *data base* yakni *forward*, *turn left*, *turn right*, *stop*, dan *shutdown now*. Selanjutnya sistem membandingkan suara masukan manusia tersebut terhadap suara pada *data base*, dan mengambil keputusan data mana yang sesuai dengan masukan suara pengguna tersebut.

Motor pada robot kursi roda bergerak sesuai dengan perintah yang telah diterima sebelumnya dari pengguna. Sistem pada robot kursi roda terus menunggu masukan dari pengguna sampai sistem pada robot tersebut dimatikan.

3.1 Blok Diagram Sistem

Blok diagram sistem ditunjukkan pada gambar 1

Gambar 1 Blok diagram sistem

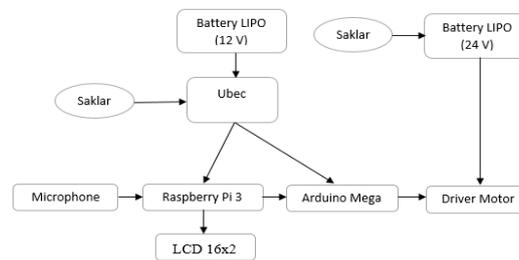


3.2 Perancangan Perangkat Keras

Pada perancangan perangkat keras, penulis mengklasifikasikannya menjadi dua perancangan, yaitu elektronika dan mekanik. Pada perancangan elektronika dibahas mengenai sistem elektronika yang dipakai

dan aplikasinya pada sistem yang telah dibuat. Pada perancangan mekanik akan dibahas mengenai tata letak perangkat dan bahan yang digunakan dalam pembuatan sistem.

3.3 Perancangan Elektronik



Gambar 2 Blok diagram elektronika

3.3.1 Catu Daya

Catu daya yang digunakan pada robot kursi roda adalah battery Li-Po 11.1 V 2200 mAh serta baterai Li-Po 22.2 V 5000 mAh. Battery Li-Po 11.1 V 2200 mAh difungsikan untuk mencatu Raspberry Pi, Arduino serta *driver* motor EMS 30A, sedangkan baterai 22.2V difungsikan untuk mencatu *driver* motor.

3.3.2 Ubec

Ubec pada robot kursi roda digunakan sebagai regulator tegangan dengan keluaran tegangan 5V disesuaikan dengan kebutuhan untuk mencatu Raspberry Pi yang membutuhkan tegangan masukan 5V untuk beroperasi dengan baik, serta digunakan untuk mencatu *driver* motor yang membutuhkan tegangan 5V agar tidak membebani Arduino apabila mengambil catuan darinya.

3.3.3 USB Mikrofon

Pada robot kursi roda, USB Mikrofon difungsikan untuk menangkap sinyal suara penggunanya. Sinyal suara tersebut dikirim ke Raspberry Pi untuk diproses sehingga mendapatkan hasil berupa perintah untuk menggerakkan robot kursi roda. USB mikrofon yang digunakan memiliki rentang frekuensi yang diterima antara 20Hz – 16KHz serta tegangan kerja 5V.

3.3.4 Raspberry Pi 3

Pada robot kursi roda, Raspberry Pi berfungsi untuk mengolah sinyal suara yang berasal dari pengguna. Hasil pengolahan sinyal tersebut ditampilkan pada LCD serta mengirimkannya kepada Arduino untuk mengatur arah gerak robot kursi roda. Untuk mengirim data hasil pengolahan sinyal suara pada Raspberry Pi, digunakan pin 2 dan pin 29 pada Raspberry Pi yang dihubungkan dengan pin 44 dan pin 46 pada Arduino. Setiap perintah yang ada pada Raspberry Pi mempunyai aksi berbeda terhadap tegangan keluaran pada pin 2 dan pin 29, kondisi tersebut dibuat agar Arduino mengetahui perintah apa yang dikirim oleh raspberry pi

3.3.5 Arduino Mega 2560

Pada robot kursi roda Arduino berfungsi untuk mengontrol arah gerak robot, dimana Arduino terhubung dengan Raspberry Pi untuk mendapatkan informasi dari keluaran Raspberry Pi berupa tegangan pada pin yang terhubung. Setiap data yang diterima oleh Arduino diproses dan hasilnya berupa perintah yang dikirim kepada *driver* motor untuk mengatur arah dan kecepatan motor.

3.3.6 EMS 30A

EMS 30A pada robot kursi roda berfungsi untuk mengatur arah gerak dan kecepatan motor berdasarkan perintah dari Arduino. Untuk menjalankan fungsinya dengan baik EMS 30 A memerlukan *trigger* berupa tegangan 24 V, tegangan 5 V serta nilai direksi dan PWM yang didapat dari Arduino.

3.3.7 LCD 16x2

LCD pada robot kursi roda difungsikan untuk menampilkan kondisi sistem *speech recognition* pada Raspberry Pi agar pengguna dapat dengan mudah mengetahui perintahnya sesuai atau tidak.

3.4 Perancangan Mekanik

3.4.1 Perancangan Aktuator Robot

Pada perancangan aktuator robot, penulis membahas penempatan motor serta menghubungkannya dengan roda pada kursi roda agar dapat berfungsi sebagai penggerak.

3.4.2 Perancangan Tempat Elektronika

Elektronika robot kursi roda diletakkan pada bagian sebelah kiri dan menggantung pada robot, adapun alasan peletakkan kotak elektronika tersebut adalah agar mikrofon dapat menjangkau pengguna serta pengguna dapat lebih mudah memonitoring kondisi dari sistem *Speech Recognition* dikarenakan pada kotak elektronika tersebut terdapat LCD yang menampilkan status dari sistem tersebut.

3.5 Perancangan Perangkat Lunak

Pada perancangan perangkat lunak penulis menggunakan algoritma MFCC untuk mengekstrak sinyal suara manusia yang masuk. Setelah mendapat hasil ekstrak sinyal yang masuk, dilanjutkan dengan membandingkan masukan sebelumnya dengan database yang terdapat pada *SD card Raspberry Pi*. Pada kasus ini database suara sudah dimasukkan kedalam sebuah format tertentu yang hanya dapat dibaca oleh *library CMU Sphinx*. Setelah sistem berhasil mengenali suara yang dimaksud pengguna, sistem kemudian mengambil data keluaran untuk menggerakkan robot kursi roda yang sesuai dengan perintah yang diinginkan dari *SD card*.

Implementasikan ke alat menggunakan *library CMU Sphinx*. *CMU Sphinx* menggunakan metode ekstraksi ciri MFCC dengan jumlah filter bank 31, jumlah koefisien MFCC 13 dan FFT sebanyak 256. Pemilihan *CMU Sphinx* dilakukan karena *library* ini jauh lebih ringan dalam hal komputasi dibanding Matlab. Mengingat processor yang di gunakan adalah *Raspberry Pi* dimana kemampuannya tidak sebaik *Personal Computer (PC)* maka pemilihan *CMU Sphinx* di harapkan dapat mengurangi beban komputasi pada CPU.

4. Pengujian dan Analisa

Pada bab ini dibahas mengenai pengujian dan analisa terhadap realisasi alat berdasarkan perencanaan dari sistem yang telah dibuat pada bab sebelumnya. Pengujian ini meliputi pengujian sistem pada kondisi tanpa kursi roda serta kondisi saat penggunaan robot kursi roda secara langsung. Pengujian ini bertujuan untuk mengetahui kualitas dari sistem yang telah dirancang.

4.1 Pengujian Tanpa Kursi Roda

Pengujian ini dilakukan tanpa menggabungkan elektronika dengan kursi roda serta dilakukan pada kondisi yang sepi. Tujuan dari pengujian ini adalah untuk mengetahui kualitas sistem *speech recognition* yang telah dibuat sebelum digabungkan dengan kursi roda, agar mendapatkan kesimpulan kelayakan sistem tersebut untuk digunakan pada alat kursi roda otomatis. Adapun beberapa pengujian yang dilakukan untuk mengetahui kualitas dari sistem yang telah dibuat.

4.1.1 Pengujian Waktu Booting Raspberry Pi

Pengujian ini bertujuan untuk mengetahui waktu yang dibutuhkan Raspberry Pi masuk ke program utama *speech recognition*.

Table 4.1 hasil pengujian waktu booting raspberry pi

Pengujian ke-	Waktu (Detik)
1	46.52
2	43.76
3	46.85
4	43.72
5	43.90
6	42.26
7	42.3
8	43.67
9	42.85
10	43.05

Dari 10 kali pengujian yang telah dilakukan, didapat rata-rata waktu Raspberry Pi untuk masuk ke program *speech recognition* adalah 43.888 detik

4.1.2 Pengujian Respon Waktu Sistem Untuk Melakukan Pengenalan Ucapan

Pengujian ini bertujuan untuk mengetahui waktu yang dibutuhkan sistem untuk melakukan pemrosesan suara sampai pengambilan keputusan.

Table 4.2 hasil pengujian waktu sistem melakukan pengenalan ucapan

Pengujian ke-	Waktu (Detik)	Pengujian ke-	Waktu (Detik)
1	6.67	21	8.81
2	4.94	22	8.04
3	4.42	23	9.23
4	4.58	24	9.16
5	5.23	25	8.33
6	4.01	26	8.32
7	3.79	27	7.64
8	4.52	28	5.46

9	4.03	29	6.24
10	4.21	30	4.66
11	4.86	31	4.69
12	5.7	32	7.01
13	6.84	33	6.9
14	5.32	34	4.31
15	4.31	35	4.25
16	4.85	36	4.55
17	4.6	37	4.44
18	5.48	38	4.5
19	4.45	39	4.38
20	5.58	40	4.56

Dari 40 kali pengujian yang telah dilakukan, didapat rata-rata waktu yang dibutuhkan sistem untuk melakukan pemrosesan suara sampai pengambilan keputusan adalah 5.59 detik. Adapun yang mempengaruhi waktu respon sistem untuk melakukan pengenalan suara yakni adanya *noise* sekitar sehingga sistem terus melakukan perekaman suara. Selain itu perangkat yang memiliki spesifikasi kurang mencukupi juga menjadi salah satu penyebab lambatnya respon sistem untuk mengenali ucapan.

4.1.3 Pengujian Pengaruh Jarak Mikrophone Terhadap Ketepatan Pengenalan Ucapan

Pengujian ini bertujuan untuk mengetahui pengaruh jarak microphone terhadap ketepatan pengenalan ucapan pada sistem.

Table 4.3 hasil pengujian pengaruh jarak microphone

Jarak Mikrophone	Percobaan ke-									
	1	2	3	4	5	6	7	8	9	10
10 cm	V	V	V	V	V	V	V	V	V	V
20 cm	V	V	V	V	V	V	X	V	V	V
40 cm	V	X	X	V	V	V	X	V	V	X
60 cm	V	X	X	V	X	X	V	V	X	X

Keterangan simbol : V = Benar
X = Salah

Dari 10 kali pengujian yang telah dilakukan, dengan menggunakan 4 jarak yang berbeda disimpulkan bahwa jarak 10 cm pada microphone merupakan jarak terbaik untuk sistem pengenalan suara.

4.1.4 Pengujian Pengaruh Frame Rate terhadap waktu dan kualitas pengenalan suara

Tujuan pengujian ini adalah untuk mengetahui pengaruh Frame rate terhadap kualitas pengenalan suara. Pengujian ini dilakukan di ruangan yang sunyi untuk mengurangi noise pada lingkungan sekitar. Adapun pengujian dilakukan dengan cara mengubah parameter pada CMU Sphinx dari 16000 menjadi 32000 serta melakukan pengujian di tempat ruangan yang sunyi.

Tabel 4. 1 Tabel pengaruh *frame rate*

Pengujian ke-	Waktu (detik)	Ketepatan pengenalan
1	6.5	√
2	5.46	√
3	5.62	√
4	7.48	√
5	7.71	√
6	7.36	√
7	6.36	√
8	5.66	√
9	5.58	√
10	5.76	√
11	5.62	√
12	5.1	√

13	5.85	√
14	5.56	√
15	5.91	X
16	5.45	X
17	5.71	√
18	5.6	X
19	6.6	√
20	8.1	√

Dari hasil tersebut didapatkan hasil rata-rata waktu pengenalan suara dengan frame rate 32000 adalah 6.1495 detik. adapun hasil tersebut didapat dari rumus :

$$\text{Waktu rata - rata} = \frac{\sum_{i=1}^{10} \text{waktu}}{10} \quad (4.1)$$

Sedangkan untuk kualitas pengenalan suaranya didapat persentase 85%, nilai itu didapat dari rumu :

$$\text{persentase keberhasilan} = \frac{\text{kata benar}}{\text{jumlah pengujian}} \times 100\% \quad (4.2)$$

4.2 Pengujian Dengan Kursi Roda

Tujuan dari pengujian ini adalah untuk mengetahui kualitas sistem yang sebelumnya telah dibuat diimplementasikan pada kursi roda. Pengujian ini dilakukan di ruangan yang sunyi untuk mengurangi *noise* pada lingkungan sekitar. Adapun pengujian dilakukan dengan menghitung waktu pengenalan suara dengan *stopwatch* serta melihat respon pengenalan suara.

Tabel 4. 2 Tabel Pengujian dengan kursi roda

Pengujian ke-	Waktu (detik)	Pengenalan suara
1	5.71	√
2	20.58	X
3	8.57	X
4	18.34	√
5	7.95	√
6	7.59	√
7	46.1	X
8	14.55	√
9	10.59	√
10	42.41	√
11	25.14	X
12	4.79	√
13	9.38	√
14	7.78	√
15	14.7	X
16	29.69	√
17	5.56	√
18	5.11	√
19	6.32	√
20	7.31	√

Keterangan :

√ : hasil pengenalan suara sesuai dengan kata yang diucapkan

X : hasil pengenalan suara tidak sesuai dengan kata yang diucapkan

Dari data tersebut didapat rata-rata waktu untuk pengenalan suara adalah 14.9 detik, adapun nilai tersebut didapat dari rumus :

$$\text{Waktu rata - rata} = \frac{\sum_{i=1}^{10} \text{waktu}}{10} \quad (4.3)$$

Sedangkan untuk kualitas pengenalan suaranya didapat prosentase 75%, nilai tersebut didapat dari rumus :

$$\text{persentase keberhasilan} = \frac{\text{kata benar}}{\text{jumlah pengujian}} \times 100\% \quad (4.4)$$

Adapun analisis mengenai lebih buruk nya kualitas sistem saat menggunakan kursi roda dibandingkan dengan tanpa menggunakan kursi roda adalah dikarenakan dipengaruhi lingkungan sekitarnya, saat pengujian ada suara yang dihasilkan dari gesekan Antara *gear* dan rantai serata bunyi dari motor DC yang digunakan

menyebabkan sistem terus merekam suara sehingga menyebabkan lamanya waktu untuk melakukan pengenalan suara serta mengganggu hasil dari pengenalan suara.

5. Kesimpulan dan Saran

5.1 Kesimpulan

Dari penelitian yang telah dilakukan tentang sistem pengenalan suara pada robot kursi roda untuk mengatur arah gerak robot kursi roda, didapat beberapa kesimpulan sebagai berikut :

1. Dari hasil pengujian waktu *booting* didapatkan hasil rata-rata untuk masuk ke program *speech recognition* yaitu 43.888 detik
2. Dari hasil pengujian waktu pengenalan suara oleh sistem didapatkan waktu rata-rata 5.59 detik.
3. Dari beberapa percobaan jarak mikrofon terhadap pusat suara didapatkan kesimpulan bahwa jarak 10 cm terhadap pusat suara memiliki tingkat akurasi yang lebih tinggi dibandingkan jarak 20 cm, 40 cm, dan 60 cm.
4. Dari hasil percobaan didapat bahwa waktu pengenalan suara dengan *frame rate* 16000 lebih baik dibandingkan dengan *frame rate* 32000 yakni 5.59 detik untuk *frame rate* 16000 dan 6.1945 detik untuk *frame rate* 32000.
5. Dari hasil percobaan didapatkan bahwa CMU Sphinx dapat digunakan pada sistem robot kursi roda.

5.2 Saran

1. menggunakan metode ekstraksi ciri yang lain dan kemudian dibandingkan hasilnya dengan metode MFCC.
2. melakukan preprocessing terlebih dahulu agar mengurangi tingkat error yang dikarenakan adanya noise .
3. mengganti single Board Computer dengan jenis yang lain agar mengurangi waktu booting dan waktu komputasi.
4. Mengganti jenis motor DC yang digunakan, agar mengurangi bunyi lain diluar suara pengguna

DAFTAR PUSTAKA

- [1] Rhomanzah, Donny.2015.Sistem Kecerdasan Buatan Untuk Robot Asisten Berbasis Algoritma Case Base Reasoning.Bandung:Telkom University.
- [2] Sidiq, Muslim. Desain dan Implementasi Speech Command Menggunakan Metode MFCC dan HMMs.Bandung:Universitas Telkom.
- [3] Khairunnisa, Dhita.2011.Analisa Dialek Suara Teleponi dengan Mel-Frequency dan K-Nearest Neighbor Berbasis Pengolahan Sinyal Digital.Bandung:Institut Teknologi Telkom.
- [4] NN.(2015), Training Acoustic Model For CMU Sphinx, <http://cmusphinx.sourceforge.net/wiki/tutorialam>
- [5] Ginting, Diego Perananta.2016. Perancangan Dan Implementasi Speech Recognition Untuk Mengubah Nada Dasar Instrumen Gong Pada Keteng-Keteng Elektronik. Bandung:Universitas Telkom.
- [6] Indarwati, Indri., Expose Data Penyandang Cacat Berdasarkan Klasifikasi ICF Tahun 2009. www.kemosos.go.id. Diakses pada tanggal 2 November 2016
- [7] Saftian, Hafizh. 2015. Perancangan dan Implementasi Sistem Robot Penggenggam Benda Menggunakan Fuzzy Logic. Bandung : Universitas Telkom.
- [8] Aulia, Masyithah Nur.2017. Implementasi Mel-Frequency Cepstrum Coefficients (MFCC) dan K-Nearest Neighbor (KNN) untuk Klasifikasi Ucapan Huruf Hijaiyah Bertanda Baca. Bandung:Universitas Telkom.